DOCUMENT RESUME

ED 313 783 EA 021 448

AUTHOR Michael, Price M.
TITLE Improving Personnel Evaluation Processes: A Synthesis of Research and Practice.
PUB DATE Nov 89
NOTE 21p.; Paper presented at the Annual Meeting of the Southern Regional Council on Educational Administration (Columbia, SC, November 11-14, 1989).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Bias; Elementary Secondary Education; *Evaluation Methods; *Job Performance; *Observation; *Personnel Evaluation; School Districts; Social Services; *Validity
IDENTIFIERS *Objectivity

ABSTRACT

School systems and social services agencies continue to grapple with personnel evaluation problems. One way to make personnel evaluation more useful is to improve the validity of evaluation judgments in appraising work performance. This paper proposes to: (1) review literature providing insight into evaluators' reasons for making inaccurate judgments concerning others' work performance; (2) describe common evaluation errors; and (3) provide guidelines for making evaluators' assessments more valid. A literature review identifies four major impediments to obtaining valid performance evaluations: cognitive processes (categorization and memory); innate forces (desire to be accepted, concern with self-protection and success, and affiliation needs); evaluators' personal characteristics; and limited observational opportunities. Errors commonly made by evaluators include the halo effect, evaluation bias and inconsistency, first impression or primacy error, performance dimension order, spillover effect, reliance on similarities, central tendency, and negativity effect. One approach to improving the personnel evaluation process is through training sessions. Better assessments can be obtained by coupling peer evaluations with evaluators' assessments or by using evaluation teams. Evaluator behavior affects the way the persons evaluated perceive assessment outcomes. The more frequently evaluators use observations and conferences in their appraisals, the harder it is to deny variability in personal performance or evaluator ability to judge competence. Developing research-based criteria can help reduce district biases. Evaluation systems that involve evaluators and evaluatees in the design process enhance the credibility and perceived validity of the evaluation process. (54 references) (MLH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED313783

BEST COPY AVAILABLE

Improving Personnel Evaluation Processes:

A Synthesis of Research and Practice

Price M. Michael, Ed.D.
Professor of Education
School of Education
West Georgia College
Carrollton, GA 30118

This paper prepared for the Annual Conference of Southern Regional Council on Educational Administration. Held November 12-14, 1989, Columbia, South Carolina.

EA 021 445

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Price M. Michael

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Improving Personnel Evaluation Processes:

A Synthesis of Research and Practice

Maintaining and strengthening desirable performance by employees which enable organizations to achieve desired goals are key contributions of personnel evaluation processes.

We are all aware that performance evaluation is an important part of an administrator's responsibility. The results of performance evaluations are used for decisions related to promotion, wage and salary administration, training, staff development, termination, etc. However, performance evaluation is one of the most problematic areas in personnel administration.

Presently, there is a move by states for mandated annual performance evaluations of school personnel using standardized instruments designed by or under the supervision of state departments of education. A requirement that the evaluators be trained to use the instruments is a part of this evaluation process.

Principals and other leadership personnel are being "trained" to use these performance evaluation instruments. These training sessions are commonly held for 3-6 days and focus on special terms, constructed dimensions, descriptors, and gaining consistency of scoring observed teaching behaviors.

Primarily, the instruments for performance evaluation are viewed as helpful in communicating a common language of expectations. The evaluation process focuses the attention of the evaluator on what is to be observed, and (with training of the evaluator) the instruments scoring processes can permit consistency in performance ratings which may result in greater evaluation validity. The preconference permitted by some of the evaluation procedures and required of others, contributes to some forms of joint goal setting.

Some researchers suggest that one way to increase the acceptance of evaluation is through participation in goal setting. These researchers argue that participation in goal setting is fundamental in reaching consensus, quickening participants' support and commitment to a process, or effecting change (Earley, P.C., 1987; Wynn & Guditus, 1984).

Educational Administration and Supervision preparation programs must play a greater role in preparing leaders as performance evaluators. Professors preparing educational leaders have a responsibility beyond the "training" on the instrument by state departments of education and state approved regional and local systems trainers.

We suggest in this paper that one way to make personnel evaluation more useful for all involved is by improving the validity of evaluation judgments in appraising work performance.

Continuous attempts have been made to address the issue of validity of performance evaluations since they became a formalized process in the early 1900s in United States' institutions. Federal, state, and local governments have made multiple changes in their employee appraisal systems over the last 80 years in order to achieve a greater degree of accuracy.

School systems and social services agencies, continue to grapple with the evaluation problem. It became more pronounced with the advent of accountability mandates from all levels of government for programs receiving funding in those agencies. School systems are increasingly being observed for indications of improved performance. "Better Schools" and "Career Advancement" programs of various types are being created by state governments and require valid evaluator assessments.

Purpose of Paper

The purpose of this paper is threefold: (a) to review literature that provides insight as to reasons evaluators make inaccurate judgments about the work performance of others, (b) to describe the common errors made by evaluators in evaluating employees, (c) to provide guidelines for making evaluators' assessments more valid. It is believed that information of this nature is of particular interest to people responsible for leadership preparation programs, and to administrators in school systems and other human service organizations, many of whom place great reliance on the judgments of supervisors in evaluating performance.

BARRIERS TO VALID PERFORMANCE EVALUATIONS

A review of the literature identifies four major aspects related to impediments to obtaining valid performance evaluations: cognitive processes, innate type forces, personal characteristics of evaluators and observational opportunities. With the exception of personal characteristics, it would appear that many of the behaviors that emanate from these factors are unconsciously or subconsciously derived.

Cognitive Processes as Barriers

Various cognitive processes of evaluators have been frequently cited in appraisal literature as obstacles to producing accurate, unbiased and reliable judgments of evaluatees. Ilgen and Feldman (1983) theorize the evaluators' task as a memory-based judgment process that entails what one has attended to in the past, the way in which information is structured and stored in memory, and other factors influencing the recall of information. Landy and Farr (1980) make a major contribution to this issue.

Five cognitive processes are identified as having the potential to distort evaluators' judgments; the first being attention. This process involves information about a person being brought to mind through automatic and controlled processes. The automatic process results in one gathering information about another with a minimal amount of time and energy being expended. Attention to the work attributes of a person occurs with very little awareness, and frequently irrelevant and inaccurate information is processed. The controlled process in attention is influenced by such factors as the uniqueness of one being evaluated, i.e., the only black person in a work group. First impression bias that infiltrates performance assessments can be explained by the nature of these processes of attention.

Categorization and memory - In this phase, the amount of information to be processed is reduced and stored by matching the features of a person and/or behavior with the features of an evaluator's prototype. A categorical assignment is made when a "goodness of fit" between an evaluatee and an evaluator's prototype is made. Thus, in

performance evaluations, opinions about an individual are derived based on cognitive representations of a category and not the unique, relevant job behavior features of the evaluatee. This process is the same as stereotyping because the features of the group's prototype become characteristic of the individual. Categorized information then enters short-term memory banks and some of the information enters long-term memory banks where it may be stored indefinitely.

Recall - Information is stored in the order of receipt; therefore, the most recent information is the most accessible and most easily brought to a level of consciousness. Superior work performed close to the time of one's evaluation would possibly result in higher ratings but the information recalled would not be a valid overall assessment of job performance.

Information search occurs when additional information is sought because contradictory data have been obtained about an evaluatee. This process of searching for information can be detrimental to evaluations when evaluators tend to seek information about or notice only those behaviors that confirm their earlier beliefs instead of noting disconfirming behaviors to originally held opinions.

Recategorization results in people changing their judgments about others because new information received is so extremely different that it cannot be disregarded. Memories that match the new category become more available in one's mind. The new information is highly negative or positive and may not represent a valid measure of work performance.

Writers who address various cognitive processes as barriers are Tverski (1977), attention; Behling, Gilford and Tolliner (1980), Wyer and Srull (1980), categorization and memory; Dreben, Fiske and Hastie (1979), recall; Snyder and Swan (1978), information search; and Snyder and Uranowitz (1978), recategorization.

Innate Forces

Henderson (1984) identified certain forces thought to be deep within the genetic and learned makeup of all people that can negatively affect the behavior of evaluators.

Desire to be accepted - All humans have a need for social affiliation and require a certain degree of social interaction. In judging another's work, if it is believed a true assessment would damage the interactive process, there will be a tendency to make the assessment more positive, thereby producing an invalid evaluation.

Central office respondents believed that the conflict between the principal as instructional leader and evaluator has not been settled, noting that collegial relationships lead many evaluators to want to be "good guys". Many felt that evaluator evaluations were upwardly biased (Wise et al. 1984). Evaluators may want to overrate teachers so as to preserve harmonious working relations in the school (Wise et al. 1984). Buser and Pace (1988) suggest that an evaluator's popularity has a tendency to be directly related to the evaluation of the evaluatee--the higher the ratings, the greater the popularity with personnel.

Concern with self-protection - People tend not to seek out experiences that will put them in danger. Concern that an evaluatee might react violently to a negative evaluation could result in inaccurate performance ratings.

Political support from "downtown" is crucial -- sometimes evaluators will not act on observed personnel problems because they are afraid that those "downtown" in positions of authority will not back them up if a probation placement becomes a heated issue (McLaughlin 1984). Evaluators must feel confident that in the instance of a warranted probation those "downtown" will back them up (Wise et al. 1984) as cited by McLaughlin (1984) in his study.

Affiliation with those holding similar views or having similar qualities - It is easier to relate to persons who are similar. People tend to react negatively to persons or situations that are dissimilar.

Concern with success - People generally seek to put themselves in situations where they will not fail. For example, supervisors as evaluators realize that their success depends on the cooperation of subordinates. This influencing force could result in a poor performance in a work group receiving a more positive appraisal than warranted in order to enhance other's perceptions of the group's success.

Personal Characteristics of Evaluators

Three job-related variables of evaluators that are thought to impact evaluations are job experience, performance level and leadership style. Researchers who have studied the influence of job experience report mixed results as to its effect. Mandell (1956) found that evaluators who had more than four years of supervisory experience tended to be more lenient in their judgments of subordinates' work; however, no significant effect was found in Klores' (1966) study. It is believed that the length of job experience could influence one's knowledge of the evaluatee and the job being performed.

Evaluator performance level has not been recently reported in research studies, although earlier works of Schneider & Bayroff (1953) and Kirchner & Reisberg (1965) did present evidence indicating that high performers were more valid in predicting job performance and that they made less rating errors.

Bernardin & Buckley (1981) and earlier researchers on supervisor's leadership styles reported similar results: evaluators described by subordinates as lower in consideration were more apt to give unfavorable ratings.

The effect of certain demographic characteristics on personnel ratings has received considerable attention in research studies. In gender studies, which have been conducted mainly in simulated work settings, females tended to give higher ratings than males and higher ratings to males than to females on some dimensions of performance (London & Poplawski, 1976; Hammer, Kim, Baird & Bigoness, 1974).

Studies on the effects of age on evaluation have generally found no significant relationship (Cleveland & Landy, 1981; Barnes, 1980; Schwab & Heneman, 1978).

Observational Opportunities as Barriers

Lopez (1968) presents four factors that influence the observation of performance: (1) position of the observer, (2) frequency of observations, (3) predisposition of the observer and (4) random error. These factors are very pertinent to evaluators of school and human service agency workers who are professional, perform their tasks in an

autonomous manner and are not accepting of or in need of direct and frequent observations on the part of a supervisor. If the predisposition of the observer (evaluator), is skewed positively or negatively toward the evaluatee the observations become distorted, the rationale already having been cited as being related to cognitive processes.

Henderson (1984) provides additional thoughts related to barriers due to opportunities to observe work performance. He hypothesizes that the fewer opportunities an evaluator has to observe work performance, the more dependent the person will be on inference and memory in identifying past performance. He also brings to mind the strong possibility that an employee's behavior will change if he/she is aware of being observed, thus the behavior may not be typical of usual job performance.

Ilgen & Feldman (1983) present the issue of limited observational opportunities that are caused by a lack of synchronization between the nature of the job requirements and the time of the appraisal. They found that evaluation policies of companies are often set arbitrarily without considering an evaluatee's work cycle.

ERRORS OF EVALUATORS IN PERFORMANCE EVALUATIONS

It is evident, based on the prior section, that errors will be made by evaluators in assessing job performance. Most articles on this topic have considered the errors in relation to performance ratings made on scales. It is our belief that these errors continue with the newly developed "effective practices" evaluation instruments that are being used in school systems in several states. Errors commonly cited as being frequently made by evaluators are listed below.

1. **The halo effect**, which entails scoring an evaluatee as superior in one quality, which influences giving a higher than deserved score on other qualities or work behaviors (Cooper, 1981).

Evaluatees report that rating scales used by their evaluators often did not deal with teaching-learning variables but with items such as teacher appearance, and sense of humor which could not provide accurate data to judge their teaching performance (Wise et al. 1984). Fisher

reports that a study of school districts in New Mexico revealed that only 28 percent of the criteria used by 81 percent of the districts for personnel evaluations focused on the teaching role, while 54 percent rated personal characteristics and the social, professional and membership roles of the evaluatee (Fisher, 1982).

2. **Evaluator bias and inconsistencies**, Kiley (1988) found evaluatees felt that these factors (when applying criteria during observation of evaluatee performance) cast doubt on the accuracy of the evaluation. McLaughlin suggests that recently most personnel evaluations use standard instruments completed by the evaluator after a brief observation -- evaluators usually base their observation scores on their own sense of effective practice. Although "trained", the evaluator may make a judgment and then rationalize it against the checklist criteria. Evaluation in this instance reflects the evaluators' individual preferences rather than a consistent set of criteria to inform either effectiveness or needs for improvement (McLaughlin 1984).

3. **First impression or primacy error**. The tendency for an evaluator to place a higher value on behavior or performance that occurred early in an appraisal period.

4. **Performance dimension order**. When two factors follow each other and are of a similar quality, the evaluator gives the first factor an accurate rating and rates the second in a similar manner because of their proximity on the observation form.

5. **Spillover effect**. Evaluators are influenced by past performance evaluations and give a similar appraisal for the current evaluation. However, the current job performance may not warrant that appraisal, whether good or bad.

6. **Same as me or different from me**. Evaluators give higher ratings when the evaluatee has similar qualities or characteristics and give lower ratings when there is dissimilarity between them.

7. **Central tendency**. Evaluators systematically judge evaluatees as average on all job performance behaviors. This error has been cited as the most common and most serious type of error. It can demoralize the high performing evaluatee and cause the poor performer to work at the

same or a lower level, since it is known that the judgment will always be that of satisfactory (Lopez, 1968).

8. **Negativity effect.** This error has been studied by Hamilton & Baumeister (1984) and Amabile & Glazebrook (1982). Negativity effect is said to be present when an insecure evaluator appraises an evaluatee negatively hoping the action will be interpreted by observers as a sign of high intelligence. Such behavior is based on the assumption that negative criticism is considered more intelligent and insightful than praise. Thus an evaluation might be determined by the self-presentational motives of the evaluator and not the performance of the evaluatee. An important aspect of this effect is the presence or knowledge that a third party will be observing the act of assessing or reviewing the evaluator's assessment.

IMPROVING THE VALIDITY OF EVALUATORS' APPRAISALS

We have explained what barriers cause evaluators to give invalid appraisals and how the barriers are operationalized as errors in performance evaluations. The problems identified are so varied and of such magnitude that one would wonder whether evaluators can be assisted in making more valid assessments of the job performance of evaluatees. This final section serves to summarize the information reported and provides proposals related to increasing the accuracy of evaluators.

Steps to Improving the Personnel Evaluation Processes

1. **Training sessions can be used to improve an evaluator's assessment ability.**

Most of the training program proposals have been designed to reduce rating errors. (Latham, Wexley, & Pursell, 1975; Borman, 1975; Bernardin & Buckley, 1981).

Ilgen and Feldman (1983) proposed the purpose of training be expanded from a concentration on reducing errors to an emphasis on the total appraisal process. They recommended that training be given on strategies and tactics for information gathering since evaluators often

make judgments based on inadequate information and usually seek only information that confirms their expectations. Evaluators should also be made aware of the behaviors and cognitive processes that inhibit them from assessing others accurately. Ilgen and Feldman recommend training be given related to the cause of performance. As an example, evaluators often do not realize the impact situational factors have on performance.

The role of training in assisting evaluators has also been addressed by Hyde and Smith (1982) who developed a model based on the assumption that training can be effective if it is extensive, allows for practice sessions and follows a systematic process. They proposed four objectives of training:

1. Teach evaluators to assess and develop job descriptions. This is important because the assessment of the performance of an evaluatee is usually dependent on the evaluator comparing job requirements with work behaviors. Experience should be given in writing job descriptions based on a job analysis approach; therefore, the evaluator will be familiar with the critical job elements comprising a particular job.

2. Familiarize the evaluator with the purpose of a performance evaluation system. This would entail training on why forms need to be completed and how the results of an appraisal will be used by an organization and the weaknesses and problems of the system.

3. Provide a detailed overview of errors made by evaluators and ways in which to avoid them.

4. Provide practice sessions of the evaluating process. This could include practice in observing, in scoring employees' behavior on forms, in communicating results to employees and setting work goals with individuals.

Because administrators are often assumed to be competent evaluators, they receive little training in standardizing procedures or maintaining proper competency levels (Buttram and Wilson 1987). Buser and Pace (1988) and Grier (1987) also emphasized the importance of evaluator training to the evaluation process. Many states and school districts are implementing programs for evaluation training to ensure evaluators are more reliable and fair. McLaughlin (1984) included this trend in evaluator training in his writings.

Training programs are including 3-5 work sessions, summer workshops to update evaluators' evaluation and supervisory skills such as observation skills and evaluating techniques in order that evaluations yielded more concrete information. Many training programs are providing practice in evaluations through activities such as simulations, case studies, role playing, and judging videotaped teaching sequences. Total training time is varying from thirty to forty hours depending on the district (Wickert 1987).

The Wise et al. (1984) study showed effective districts implement measures to monitor evaluator competency in judging the evaluatee performance on personnel evaluations. One such measure is by using central office personnel to review the evaluatee performance evaluations to ensure credibility and uniformity. Buttram and Wilson (1987) cite Wise et al. (1984), suggesting that evaluators should be held accountable for their judgments. McLaughlin (1984) expressed similar views on the evaluation of evaluators by superiors.

Wickert (1987) suggests a required level of administrator competence within each district in order to ensure evaluations are fair to the evaluatees. McLaughlin (1984) suggests that training provides evaluators and evaluatees with a common language. Evaluators are then able to move beyond global statements about evaluatee performance and discuss specifically about the observed evaluatee performance as well as provide concrete examples. The best supported and carefully constructed process will founder if evaluators lack the necessary background, knowledge, and expertise (Wise et al. 1984).

2. **Better assessments can be obtained if peer evaluations are coupled with evaluators', i.e. director supervisors' assessments.**

Because of the continuous opportunity to observe others in work groups, it is believed peers can give excellent insight into behaviors demonstrated by persons in their work groups. Peers can also be a source for identifying leadership skills and potential in others (Henderson, 1984).

Specific criteria must be present to ensure the accuracy of peer evaluations. Mumford (1983) suggests the following conditions:

- opportunity to observe criterion-relevant task performance
- knowledge of abilities and compentencies relevant to task performance
- strong personal incentives for evaluating peers
- similarity in work performed and job level between peers and evaluatees

3. **Evaluation teams will produce more valid appraisals than a supervisor's sole judgment.**

Teams - Team Evaluation and Management System, a model proposed by Edwards and Sproull (1985) is based on the assumption that evaluation teams will produce more valid appraisals than a supervisor's sole judgment. The employee to be evaluated selects a team of five persons comprised of an immediate supervisor and fellow employees who are within two organizational levels of the evaluatee. Peers must have adequate contact with the evaluatee in order to make informed judgments. Team members are asked to judge the performance and potential of two individuals at a time. The supervisor's evaluation accounts for 20% of the appraisal. Bias and inadequate information on the part of any one team member is minimized due to others' responses. The evaluation forms provide a record of judgments made and will indicate significant differences of individuals from the view of the majority. It is reported that organizations that have adopted the model have found it to be a satisfying process for employees as well as supervisors.

4. **Henderson (1984) has provided the following set of factors he believes improves the validity of appraisals made by evaluators:**

- evaluator has observed and is familiar with behaviors, work to be appraised
- evaluator has documented behaviors, critical incidents to improve recall
- an instrument is provided to an evaluator to obtain and review job performance-related information
- evaluator is cognizant of personal biases and is committed to taking action to minimize their effect
- attention by evaluator is focused to performance related behavior over which evaluatee has greater control
- evaluator is trained to properly use the appraisal system
- evaluator's rating score is summarized and compared with other evaluators
- higher level of management is held accountable for reviewing all appraisals.

5. **Evaluators' behavior affects the way evaluatees perceive the outcome of personnel evaluations.**

An evaluatee is more likely to accept the evaluator's recommendations for performance improvement if the evaluatee respects the judgment of the evaluator. An evaluatee is also more likely to feel more stable and supported in his/her duties if the evaluatee perceives the evaluator as having strong and consistent ethical values. Evaluators must be a model of ethical behavior to evaluate personnel performance fairly and for the evaluation to be valid (Seldin 1988).

The evaluatee should come out of the evaluation feeling whole and not assassinated--this positive feeling is one that evaluators may promote by showing both concern for performance improvement as well as respect for the evaluatee. Showing concern for an evaluatee's professional growth is of inestimable value in building credibility as

an evaluator (Fisher 1982). Laing (1986) also expressed similar views of the value and worth of personnel.

6. **The more frequently evaluators utilize observations and conferences in their appraisals, the more difficult it will be to deny variability in personnel performance or evaluator ability to judge competence (Buser and Pace 1988).**

Regular classroom observations by a trained evaluator can provide feedback concerning evaluatee strengths, weaknesses, and suggestions for improvement that are necessary to evaluatee satisfaction, efficacy, and growth. The evaluatee will be more likely to regard the evaluator as competent and view the evaluation process as a rewarding experience (McLaughlin 1984). Fisher (1982) also stressed the importance of observation. He suggests that evaluatees need immediate feedback after each observation--evaluatee strengths as well as suggestions for improving evaluatee performance. A good observation lasts two hours and that evaluator notes should be objective--reflecting what is actually happening and not the evaluator's opinion (Ellis 1985).

7. **Developing evaluation criteria based on literature on effective evaluatee practices can serve as a means of reducing the districts thoughts and biases.**

Buttram and Wilson (1987) cite Caldwell (1986), the Delaware Department of Public Instruction, and the Florida Coalition (1983) as reporting that many districts have begun developing evaluation criteria based on literature on effective evaluatee practices as a means of reducing the districts thoughts and biases influence on the evaluation criteria. Grier (1987) also states similar ideas about the effects of literature on evaluation criteria.

8. **Self-evaluation coupled with the classroom observations would then be analyzed by the evaluator along with the evaluation results to determine evaluatee strengths and suggestions for improved performance.**

Fisher (1982) reports that he has evaluatees fill out a self-evaluation, approximately four pages in length, noting what was going well and areas the evaluatees would want to improve in their own

performance and the classroom program prior to the evaluation. This self-evaluation coupled with the classroom observations would then be analyzed by the evaluator along with the evaluation results to determine evaluatee strengths and suggestions for improved performance. Wise et al. (1984) cites Bodine (1973), Bushman (1974), and Riley and Schaffer (1979) as being among the authors of teacher evaluation literature that has begun to recognize the importance of self-assessment and allowing teacher input into the determination of criteria and standards.

Evaluation systems where evaluators and evaluatees were involved in designing the system enhance the credibility and perceived validity of the evaluation process (Buser and Pace 1988). For an evaluation system to succeed, evaluatees must see it as equitable and relevant. Evaluatee participation in the evaluation design is crucial to evaluatee commitment to change based on evaluation outcomes (McLaughlin 1984). Evaluatee involvement in the evaluation process creates a sense of ownership and improves the legitimacy, focus, and meaning of the activity (Huddle 1985). When evaluatees have a say in the decisions that affect them, they feel less isolated and more a part of the school (Krajewski 1986).

9. To increase the validity of evaluators' judgments, the evaluation process requires careful documentation of teaching behaviors that lead to unsatisfactory ratings (i.e., ratings of "needs improvement").

Careful documentation enables someone other than the evaluator to verify that the teaching criteria have been applied appropriately. Valid, reliable, and helpful evaluation requires evaluators who recognize good teaching (and its absence) and who know how to improve poor teaching when they find it (Wise et al. 1984).

Evaluatee involvement and responsibility improve the quality of personnel evaluation. School districts could use teacher organizations to help design and oversee personnel evaluations to ensure its legitimacy, fairness, and effectiveness.

CONCLUSION

Most school systems and other human service organizations will agree that evaluation of personnel is necessary for providing both the evaluatee and the organization with feedback helpful for personal growth, and decisions related to human resource selection and development. Some problems associated with acceptance of this personnel function can be understood by focusing upon psychological phenomena that act as barriers to valid performance evaluations.

Working to improve the validity of performance evaluations will provide positive perceptions of the process and increase the palatability of policies implemented to carry out this personnel function.

Based upon this review of the literature and practices, the issue of the evaluator as a major factor in producing invalid performance evaluations is a significant one. The causes for an evaluator's errors seem to be primarily at an unconscious level or due to inadequate opportunities to observe and/or review the work of an evaluatee. Programs preparing educational administrators must give more attention to the psychological phenomena that act as barriers to valid performance evaluations in our training programs.

Most of the recommendations for improving the validity of an evaluator's appraisal were presented as models based on assumptions. There is a need for additional studies of the suggested solutions. The good news is some state departments of education, state and school systems are developing evaluation programs and instruments that are beginning to reflect some of the research and practices reported in this paper.

REFERENCES

Amabile, T. M. (1983). Brilliant but cruel: Perceptions of negative evaluators. Journal of Experimental Social Psychology, 14, 146-155.

Amabile, T. M. & Glazebrook, A. (1982). A negativity bias in interpersonal evaluation. Journal of Experimental Social Psychology, 18, 1-21.

Barnes, J. L. (1980). Age as a perceptual cue in performance appraisal process: The relationship between perceived age similarity and performance ratings of exempt employees. Unpublished doctoral dissertation, The Pennsylvania State University.

Behling, V. W., Gifford, W. E., & Tolliner, J. M. (1980). Effects of grouping information on decision making under risk. Decision Sciences, 11, 272-283.

Bernardin, H. J., & Beaty, R. W. (1984). Performance appraisal: Assessing human behavior at work. Boston: Kent Publishing.

Bernardin, H. J. & Buckley, M. R. (1981). A consideration of strategies in rater training. Academy of Management Review, 6, 205-212.

Borman, W. C. (October 1975). Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings. Journal of Applied Psychology, 60(5), 135-144.

Buser, R. L. and Pace, V. D. (December 1988). Personnel Evaluation: Premises, Realities, and Constraints. NASSP Bulletin, 75, 85-87.

Buttram, J. L. and Wilson, B. L. (April 1987). Promising Trends in Teacher Evaluation. Educational Leadership, 44, 4-6.

Cascio, W. F., & Valenzi, E. R. (1977). Behaviorally anchored rating scales: Effects of education and job experience of raters and ratees. Journal of Applied Psychology, 62, 278-282.

Cleveland, J. N., & Landy, F. J. (1981). The influence of rater and ratee age on two performance judgments. Personnel Psychology, 34, 19-29.

Cooper, W. H. (1981). Ubiquitious halo. Psychological Bulletin, 90, 218-244.

Dreben, E. K., Fiske, S. T., & Hastie, R. (1979). The independence of evaluative and item information: Impression and recall order of effects in behavior-based impression information. Journal of Personality and Social Psychology, 37, 1758-1768.

Earley, P.C., Wojnaroski, P. & Prest, W. (1987). Task planning and energy expended: Exploration of how goals influence performances. Journal of Applied Psychology, 72, 107-114.

Edwards, M. R., & Sproull, J. R. (March 1985). Making performance evaluations perform: The use of TEAM evaluation. Personnel, 62(3), 38-32.

Ellis, T. R. (1985). Teacher Evaluation is Hard Work--and it Should be. Principal, 64, 22-24.

Fisher, N. A. (May 1982). Evaluating Teachers: One Principal's Way. Principal, 61, 37-39.

Grier, T. B. (February 1987). Courageous teacher evaluation proves a principal's prowess. The Executive Educator, 9, 17-19.

Hamilton, J. C., & Baumeister, R. F. (1984). Biasing evaluation to appear unbiased: A self-presentational paradigm. Journal of Experimental Social Psychology, 20, 552-566.

Hammer, W. C., Kim, J. S., Baird, L., & Bigoness, W. J. (1974). Race and sex as determinants of ratings by potential employers in a simulated work sampling task. Journal of Applied Psychology, 59, 705-711.

Henderson, R. I. (1984). Practical guide to performance appraisal, Reston, VA: Prentice-Hall. 1-24, 308-317.

Huddle, G. (March 1985). Teacher Evaluation--How Important for Effective Schools? Eight Messages from Research. NASSP Bulletin, 69, 58-63.

Hyde, A., & Smith, M. (1982). Performance appraisal and training: Objectives a model for change and a note of rebuttal. Public Personnel Management Journal, 11, 358-366.

Ilgen, D. R., & Feldman, J. M. (1983). Performance appraisal: A process focus. In B. Shaw and L. L. Cummings (Eds.) Research in organizational behavior. Greenwich, CT: JAI Press, 14-197.

Kiley, M. A. (1988). Teachers' and Aministrators' Views of Evaluation--Differing Perspectives. (ERIC Document Reproduction Service No. ED 300 434).

Kirchner, W. K., & Reisburg, D. J. (1965). Differences between better and less effective supervisors in appraisal of subordinates. Personnel Psychology, 15, 295-302.

Klores, M. S. (1966). Rater bias in forced distribution ratings. Personnel Psychology, 19, 411-421.

Krajewski, R. J. (January 1986). How principals can take the agony out of teacher evaluation. The Executive Educator, 8, 19 and 33.

Laing, S. O. (November 1986). The Principal and Evaluation. NASSP Bulletin, 70, 91-93.

Landy, F. J., & Farr, J. L. (1980). Performance rating. Psychological Bulletin, 87, 72-107.

Landy, F.J., & Farr, J. L. (1983). The measurement of work performance: Methods, theory, applications. New York: Simon & Shuster.

Latham, G. P., Wexley, K. N., & Pursell, E. D. (October 1975). Training managers to minimize rating errors in the observation of behavior. Journal of Applied Psychology, 60, 550-555.

Lemley, R. E. (March 1983). Evaluation: The Key to Effective Learning. NASSP Bulletin, 67, 115-119.

London, M., & Poplawski, J. R. (1976). Effects of information on stereotype development in performance appraisal and interview context. Journal of Applied Psychology, 61, 199-205.

Lopez, F. M. (1968). Evaluating employee performance. Chicago: Public Personnel Association.

Mandell, M. M. (1956). Supervisory characteristics and ratings: A summary of recent research. Personnel Psychology, 32, 435-440.

435-440.

McLaughlin, M. W. (Fall 1984). Teacher Evaluation and School Improvement. Teachers College Record, 86, 193-205.

Mumford, M. C. (1983). Social comparison theory and the evaluation of peer evaluators: A review and some implications. Personnel Psychology, 30, 867-881.

Perry, L. T., & Barney, J. B. (Winter 1981). Performance lies are hazardous to your health. Organization Dynamics, 68-80.

Rizzo, W. A., & Frank, F. D. (1977). Influence of irrelevant cues and alternate forms of graphic rating scales on the halo effect. Personnel Psychology, 30, 408-417.

Saal, F. E., Downey, R. A., Lahey, M. A. (1980). Rating the ratings: Assesssing the psychometric quality of rating data. Psychological Bulletin, 88, 413-420.

Schmidt, F. L., & Johnson, R. H. (1973). Effect of race on peer ratings in an industrial setting. Journal of Applied Psychology, 57, 237-241.

Schneider, D. E., & Bayroff, A. G. (1953). The relationship between rater characteristics and validity of ratings. Journal of Applied Psychology, 37, 378-280.

Schwab, D. P., & Heneman, H. G. (1978). Age stereotyping in performance appraisal, Journal of Applied Psychology, 63, 573-578.

Seldin, C. A. (December 1988). Ethics, Evaluation and the Secondary Principal. NASSP Bulletin, 75, 9-11.

Snyder, M., & Swan, W. B., Jr. (1980). Hypothesize-testing processes in social interaction. Journal of Personality and Social Psychology, 39, 222-234.

Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. Journal of Personality and Social Psychology, 36(9), 941-950.

Thorton, G. C. & Zorick, S. (1980). Training to improve observer accuracy. Journal of Applied Psychology, 65, 351-354.

Tverski, A. (1977). Features of similarity. Psychological Review, 84, 327-352.

Vance, R. J., Winne, P. S., & Wright, E. S. (1983). A longitudinal examination of rater and ratee effects in performance ratings. Personnel Psychology, 609-620.

Wickert, D. M. (September 1987). Using Teacher Evaluation for Improving Instruction. Clearing House, 61, 23-24.

Wise, A. E. et al. (June 1984). Teacher Evaluation: A Study of Effective Practices. Rand Corp. (ERIC Document Reproduction Service No. ED 246 559).

Wyer, R. S., & Srull, T. K. (1980). The processing of social stimulus information: A conceptual integration. In R. Hatie, et.al., (Eds.) Person memory and encoding processes. Hillsdale, NJ: Erlbaum.

Wynn, R., & Guditus, C. (1984). Team management: Leadership by consensus. Columbus, OH: Charles E. Merrill.